

B52



PCT

WORLD INTELLECTUAL PROPERTY  
International Bu:

WO 9608570A1

INTERNATIONAL APPLICATION PUBLISHED UNDER

WO 96/08570

<p>(51) International Patent Classification 6 : <b>C12N 15/62, 15/85, 5/10, C07K 14/705, 14/16, 16/00</b></p>	<p><b>A1</b></p>	<p>(11) International Publication Number: <b>WO 96/08570</b></p>
<p>(21) International Application Number: <b>PCT/US95/11720</b> (22) International Filing Date: <b>14 September 1995 (14.09.95)</b></p>		<p>(43) International Publication Date: <b>21 March 1996 (21.03.96)</b></p>
<p>(30) Priority Data: <b>08/305,700</b> <b>14 September 1994 (14.09.94)</b> <b>US</b>  (71) Applicant: <b>FUJI IMMUNOPHARMACEUTICALS CORPORATION [US/US]; 125 Hartwell Avenue, Lexington, MA 02173 (US).</b>  (72) Inventors: <b>LO, Kin-Ming; 39 Elmwood Road, Wellesley, MA 02181 (US). SUDO, Yukio; 425 Woburn Street #37, Lexington, MA 02173 (US). GILLIES, Stephen, D.; 245 Leavitt Street, Hingham, MA 02043 (US).</b>  (74) Agent: <b>PITCHER, Edmund, R.; Testa, Hurwitz &amp; Thibault, High Street Tower, 125 High Street, Boston, MA 02110 (US).</b></p>		<p>(81) Designated States: <b>AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</b>  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>

(54) Title: **EXPRESSION AND EXPORT TECHNOLOGY OF PROTEINS AS IMMUNOFUSINS**

(57) Abstract

Disclosed are DNAs produced by recombinant techniques for inducing the expression and subsequent secretion of a target protein. The DNAs encode, in their 5' to 3' direction, a secretion cassette, including a signal sequence and an immunoglobulin Fc region, and a target protein. The DNAs can be transfected into a host cell for the expression, production and subsequent secretion of the target protein as a fusion protein. The secreted protein can be collected from the extracellular space, and further purified as desired. The secreted fusion protein additionally can be proteolytically cleaved to release the target protein from the secretion cassette.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

EXPRESSION AND EXPORT TECHNOLOGY  
OF PROTEINS AS IMMUNOFUSINS

This patent application is a continuation-in-part of U.S. Serial Number 08/305,700, filed on September 14, 1994.

5 Background

The invention relates to fusion protein expression systems for use in mammalian cells that enhance the production of a given target protein. More specifically, the invention relates to a secretion  
10 cassette, comprised of a mammalian signal peptide and a portion of mammalian immunoglobulins, which, when used as the amino-terminal fusion partner to the target protein, generally leads to high level expression and secretion of the fusion product. Such fusion proteins are useful, for example, for the production and extracellular  
15 collection of target proteins without the need for lysis of a host cell. The invention is perhaps most useful for the expression of target proteins which are not normally secreted from a host cell, are secreted at low levels from a host cell, or are toxic or otherwise deleterious to a host cell.

20

Expression systems employing gene fusion constructs have been used to enhance the production of proteins in bacteria. Employing a bacterial protein that is normally expressed at a very high level as the amino-terminal fusion partner of a fusion protein helps to ensure  
25 efficient transcription and translation of the message, and in some cases the secretion and solubilization of the fusion protein (Smith and Johnson (1988) Gene 67:31; Hopp et al. (1988) Biotechnology 6:1204; La Vallie et al. (1993) Biotechnology 11:187).

30 The major goal of expression of recombinant fusion proteins in mammalian cells has been to confer novel properties to the hybrid molecules, e.g., targeting of a cytokine or toxin in vivo, Fc receptor binding, complement fixation, protein A binding, increasing the half-life, and crossing the blood-brain barrier. Examples of recombinant  
35 fusion proteins produced in mammalian cells include cytokine

- 2 -

immunoconjugates (Gillies et al. (1992) Proc. Natl. Acad. Sci. USA 89:1428; Gillies et al. (1993) Bioconjugate Chemistry 4:230), immunoadhesins (Capon et al. (1989) Nature 337:525), immunotoxins (Chaudhary et al. (1989) Nature 339:394), and a nerve growth factor conjugate (Friden et al. (1993) Science 259:373). Each of the foregoing publications is incorporated herein by reference. Proteins produced in mammalian cells often do not have the solubility and secretion problems encountered in bacterial expression. The use of gene fusion constructs to enhance the production or secretion of a target protein in a mammalian system has not been explored fully.

It is the object of the invention to provide DNAs which facilitate the production and secretion of a target protein. In particular, objects of the invention are to provide novel DNAs which: facilitate efficient production and secretion of hard to express proteins, such as nuclear proteins, regulatory proteins and proteins which otherwise may be toxic to a host cell, and can be adapted to any target polypeptide of interest which can be coded for and expressed in a host organism; to provide DNA constructs for the rapid and efficient production and secretion of proteins in a variety of host cells; and to provide a method for the production, secretion and collection of genetically engineered proteins, including non-native, biosynthetic, or otherwise artificial proteins, such as proteins which have been created by rational design. Other objects of the invention are to provide DNA sequences which, when fused to a polynucleotide encoding a target protein, encode a fusion polypeptide which can be purified using common reagents and techniques, and to interpose a proteolytic cleavage site between the encoded secretion cassette and the encoded target protein such that the secretion cassette can be cleaved from the target protein and the target protein can be purified independently. Still another object is to provide a procedure which is both efficient and inexpensive.

These and other objects of the invention will be apparent from the description, drawings, and claims that follow.

- 3 -

Summary Of The Invention

The present invention features a DNA of general applicability for production and secretion of fusion proteins. The DNA comprises a secretion cassette, as the amino-terminal fusion partner, and a target protein, and is termed herein an "immunofusin". The invention provides, in its various aspects, a recombinant DNA encoding the immunofusin, and methods of producing the encoded immunofusin protein. The immunofusin is a DNA which comprises a polynucleotide encoding a secretion cassette, comprising in its 5' to 3' direction a signal sequence and an immunoglobulin Fc region, and a polynucleotide encoding a target protein fused to the 3' end of the secretion cassette. A secretion cassette of the invention, once constructed, can be fused to various target proteins. Additionally, one can optimize the sequences which regulate the expression of a secretion cassette, and hence the expression of the immunofusin. The resultant DNA can be expressed at high levels in a host cell, and the fusion protein is efficiently produced and secreted from the host cell. The secreted immunofusin can be collected from the culture media without the need for lysis of the host cell, and can be assayed for activity or purified using common reagents as desired.

The portion of the DNA encoding the signal sequence preferably encodes a peptide segment which directs the secretion of the immunofusin protein and is thereafter cleaved. As used in the specification and claims, "immunoglobulin Fc region" means the carboxyl-terminal portion of an immunoglobulin heavy chain constant region. As is known, each immunoglobulin heavy chain constant region is comprised of four or five domains. The domains are named sequentially as follows: CH1-hinge-CH2-CH3(-CH4), and the Fc region of each immunoglobulin subclass lacks at least the CH1 domain. As is apparent from a review of the DNA sequences of the immunoglobulin subclasses, the DNA sequences of the heavy chain domains have cross-homology among the immunoglobulin classes, e.g., the CH2 domain of IgG is homologous to the CH2 domain of IgA and IgD, and to the CH3 domain of IgM and IgE. The portion of the DNA encoding the immunoglobulin Fc

- 4 -

region preferably comprises at least a portion of a hinge domain, and a CH3 domain of Fc $\gamma$  or the homologous domains in any of IgA, IgD, IgE, or IgM. The immunoglobulin Fc region also preferably comprises at least a portion of the DNA encoding a hinge and a CH3 domain of Fc $\gamma$  or the homologous domains in any of IgA, IgD, IgE or IgM.

The currently preferred secretion cassette is a polynucleotide encoding, in its 5' to 3' direction, the signal sequence of an immunoglobulin light chain gene and the Fc $\gamma$ 1 region of the human immunoglobulin  $\gamma$ 1 gene. The Fc $\gamma$ 1 region of the immunoglobulin  $\gamma$ 1 gene includes at least a portion of the hinge domain and CH3 domain, or at least a portion of the hinge domain, CH2 domain and CH3 domain. The DNA encoding the secretion cassette can be in its genomic configuration or its cDNA configuration. However, the studies described below use a secretion cassette in the genomic configuration. The use of human Fc $\gamma$ 1 as the Fc region sequence has several advantages. For example, if the fusion protein is to be used as a biopharmaceutical, the Fc $\gamma$ 1 domain may confer the effector function activities to the fusion protein. The effector function activities include the biological activities such as complement fixation, antibody-directed cellular cytotoxicity, ability for placental transfer, and a longer serum half-life. The Fc domain also provides for detection by anti-Fc ELISA and purification through binding to Staphylococcus aureus protein A ("Protein A"). In certain applications it may be desirable to delete specific effector functions from the Fc region, such as Fc receptor binding or complement fixation.

In another embodiment the Fc region can be a murine immunoglobulin gene. The use of murine Fc as the Fc region can have advantages. For example, if the fusion protein is to be used for the preparation of proteins in mice, then the murine Fc region will not elicit an immune response in the host animal. The Fc domain may confer the effector function activities to the fusion protein, and allow for detection of the fusion protein by anti-Fc ELISA and purification through binding to Protein A. In certain applications it may be desirable to delete specific effector functions from the Fc region.

- 5 -

In another embodiment the DNA sequence encodes a proteolytic cleavage site interposed between the secretion cassette and the target protein. A cleavage site provides for the proteolytic cleavage of the encoded fusion protein thus separating the Fc domain from the target protein. As used herein, "proteolytic cleavage site" is understood to mean the amino acid sequences which are cleaved by a proteolytic enzyme or other proteolytic cleavage agents. As will be described in more detail below, useful proteolytic cleavage sites include amino acid sequences which are recognized by proteolytic enzymes such as trypsin, plasmin or enterokinase K.

In a preferred embodiment the target protein sequence encodes prostate specific membrane antigen, PSMA. PSMA is a type II membrane protein, thus the extracellular domain, or soluble form of the protein, is utilized as the target protein sequence. The encoded soluble form of PSMA can be a human sequence such as the sequence provided in Israeli et al. (1993) Cancer Res., 53:227-ff.

In another preferred embodiment the target protein sequence encodes the protein gp120. The envelope protein gp120 of human immunodeficiency virus is a glycoprotein which is expressed in infected cells as a polyprotein, gp160, and then cleaved by a cellular protease to gp120 and gp41. The nucleotide sequence and amino acid sequence of gp120 is provided in Ratner et al., 1985, Nature, 313:277-ff.

In another aspect, the DNA sequence of the invention is integrated within a replicable expression vector. As used herein, "vector" is understood to mean any nucleic acid comprising a nucleotide sequence of interest and competent to be incorporated into a host cell and to be recombined with and integrated into the host cell genome, or to replicate autonomously as an episome. Such vectors include linear nucleic acids, plasmids, phagemids, cosmids and the like. A preferred expression vector is pdC, in which the transcription of the immunofusion DNA is placed under the control of the enhancer and promoter of the human cytomegalovirus. The vector pdC was derived from pdEMP, which is described in Lo et al. 1991, Biochim. Biophys. Acta 1088:712 (which

- 6 -

publication is incorporated herein by reference) as follows. The Sall-XhoI fragment containing the original enhancer and promoter sequence were replaced by the enhancer and promoter of the human cytomegalovirus by standard molecular biology techniques. The enhancer and promoter  
5 sequence of the human cytomegalovirus used was derived from nucleotides -601 to +7 of the sequence provided in Boshart et al., 1985, Cell 41:521, which is incorporated herein by reference. The vector also contains the mutant dihydrofolate reductase gene as a selection marker (Simonsen and Levinson (1983) Proc. Nat. Acad. Sci. USA 80:2495,  
10 incorporated herein by reference).

An appropriate host cell can be transformed or transfected with the DNA sequence of the invention, and utilized for the expression and secretion of a target protein. Currently preferred host cells for use  
15 in the invention include immortal hybridoma cells, myeloma cells, 293 cells, Chinese hamster ovary cells, Hela cells, and COS cells. As used herein, "gene expression" or "expression of a target protein" is understood to refer to the transcription of the DNA sequence, translation of the mRNA transcript, and secretion of the fusion protein  
20 product.

The method of the invention involves providing a DNA sequence encoding an immunofusin, transfecting the DNA sequence into a host cell by an available transfection or transformation technique, culturing the  
25 transfected host cell in a suitable media under conditions which promote the expression and secretion of the immunofusin, and collecting the fusion protein from the extracellular media. When desired, the target protein may be cleaved from the secretion cassette either before or after it is collected from the extracellular media.  
30

Other advantages and features of the invention will be apparent from the description, drawings, and claims which follow.



- 7 -

Brief Description of the Drawing

The Figures 1A-D are a schematic illustration of an immunofusin. Figure 1A, "DNA," illustrates the DNA encoding an immunofusin protein. Figure 1B, "Fused Protein 1," illustrates the immunofusin protein prior to cleavage of the signal sequence. Figure 1C, "Fused Protein 2," illustrates the immunofusin protein after cleavage of the signal sequence. Figure 1D, "Target Protein," illustrates the target protein portion of an immunofusin protein after cleavage of the immunofusin protein at the cleavage site which is interposed between the Fc region and the target protein.

- 8 -

Detailed Description

The present invention is a DNA comprising a polynucleotide encoding, in the 5' to 3' direction, a signal sequence, an Fc region of an immunoglobulin, and a target protein. This approach to the expression and subsequent secretion of a target protein is superior to the existing techniques because of the choice and the configuration of the secretion cassette which is placed at the 5' end of the fusion construct. Additionally, the regulatory sequences which direct the expression of the secretion cassette can be optimized, and the optimized secretion cassette can be paired with numerous target proteins, thus allowing for the efficient production of numerous fusion proteins.

The production of the immunofusion proteins is characterized as efficient and high level, because the target protein has been produced at the level of several micrograms/milliliter using the DNAs and methods according to the invention. Previously, workers in the art have rarely quantitated the expression levels of hard to express proteins due to the low levels of expression that are obtained in the known mammalian expression systems and the difficulties faced in quantitating proteins by techniques such as western blotting and RIA. Prior to the teachings of this invention, expression of microgram per milliliter of hard to express proteins would often be attempted using bacterial expression systems.

This invention is based on the concept that the ease of production and collection of a target protein could be improved if the polypeptide of interest were linked to an immunoglobulin Fc domain and the fusion protein were expressed in a host cell, in particular a complementary host cell which naturally expresses the immunoglobulin, such that the fusion protein would be readily secreted from the host cell. In addition to promoting the secretion of the fusion protein from the host cell, the Fc region can further be exploited to aid in the purification of the fused polypeptide. The general approach of the invention involves the construction of recombinant DNA which encodes a fused

- 9 -

polypeptide, which upon expression, results in expression of a secretion cassette linked to a target protein, i.e., a protein of interest having potential or demonstrable utility.

5 The overall structure of the preferred DNA of the invention, the fusion protein it encodes, the form of the protein which is most often secreted and the target protein product after enzymatic cleavage are illustrated schematically in Figures 1A-D. Reference characters in the DNA, Figure 1A, are carried over into the protein, Figures 1B-D, as  
10 corresponding primed characters. The DNA which encodes the immunofusin is shown between the start and the stop markers on the illustrated DNA sequence, Figure 1A. Upstream regulatory elements are shown at the 5' end of the DNA and are labeled "regulatory sequences". The DNA is composed of three distinct polynucleotides which are linked together.  
15 In Figure 1A, 3' of the regulatory sequences, which may be optimized for each secretion cassette, is a first DNA 8 which encodes a secretion cassette comprising two of the three polynucleotides: 1) a signal sequence 10, and 2) an immunoglobulin Fc $\gamma$  region 12. The immunoglobulin Fc $\gamma$  region is comprised of three subregions: 1) a hinge  
20 region 14, 2) a CH2 region 16, and a CH3 region 20. Attached to the 3' end of the DNA encoding the secretion cassette is the third polynucleotide, a DNA encoding the target protein 24. Optionally, DNA encoding a proteolytic cleavage site 22 can be interposed between the DNA encoding the CH3 region of the immunoglobulin Fc $\gamma$  region and the  
25 DNA encoding the target protein.

The encoded fused protein comprises the secretion cassette 8' and the target protein 24', shown as Fused Protein 1 in Fig. 1B. Most often the signal peptide 10' will be enzymatically cleaved from the  
30 fusion protein by the host cell prior to the secretion of the immunofusin, and thus Fused Protein 2, shown in Figure 1C, shows the secreted fused protein which comprises the Fc $\gamma$  peptide 12' fused to the target polypeptide 24'. Both Fused Protein 1 and Fused Protein 2 show the optional interposition of a proteolytic cleavage site 22' between  
35 the CH3 domain 20' of the Fc $\gamma$  region 12' and the target protein 24'. Cleavage of either Fused Protein with the appropriate proteolytic agent

- 10 -

at the cleavage site 22' results in the release of the target protein 24' from the Fc region 12', as shown in Figure 1D.

The processes for manipulating, amplifying and recombining DNAs are generally well known in the art, and therefore are not described in detail herein. Methods of identifying and isolating genes encoding proteins of interest, or for constructing such genes, are well understood and developed. In general the methods involve selecting genetic material coding for amino acids which define the polypeptide of interest according to the genetic code.

Accordingly, the DNA construction principle disclosed herein can be exploited using known recombinant DNA techniques involving the use of various restriction enzymes which make sequence specific cuts in DNA to produce blunt ends or cohesive ends, DNA ligase techniques enabling enzymatic addition of sticky ends to blunt ended DNA, construction of synthetic DNAs by assembly of short oligonucleotides, cDNA synthesis techniques, polymerase chain reaction, and synthetic probes for isolating genes having a particular function. Various promoter sequences and other regulatory DNA sequences used in achieving expression, and various types of host cells are also known and available. Conventional transfection techniques, and equally conventional techniques for cloning and subcloning DNA are useful in the practice of this invention and known to those skilled in the art. Various types of vectors may be used such as plasmids and viruses including animal viruses. The vectors may exploit various marker genes which impart to a successfully transfected cell a detectable phenotypic property that can be used to identify which of a family of cells has successfully incorporated the recombinant DNA of the vector. Given the foregoing state of the genetic engineering art, skilled persons are enabled to practice the invention disclosed herein in view of this disclosure.

One method for obtaining the DNA encoding the various synthetic linkers disclosed herein is by assembly of synthetic oligonucleotides in a conventional, automated, polynucleotide synthesizer followed by

- 11 -

ligation with a ligase. For example, the linkers can be synthesized as complementary DNA fragments using phosphoramidite chemistry.

The signal sequence of the invention is a polynucleotide which  
5 encodes an amino acid sequence that initiates transport of a protein  
across the membrane of the endoplasmic reticulum. Signal sequences  
which will be useful in the invention include antibody light chain  
signal sequences, e.g., antibody 14.18 (Gillies et. al., 1989, Jour. of  
Immunol. Meth., 125:191-202), antibody heavy chain signal sequences,  
10 e.g., the MOPC141 antibody heavy chain signal sequence (Sakano et al.,  
1980, Nature 286:5774), and any other signal sequences which are known  
in the art (see for example, Watson, 1984, Nucleic Acids Research  
12:5145). Each of these references is incorporated herein by  
reference. Signal sequences have been well characterized in the art  
15 and are known typically to contain 16 to 30 amino acid residues, and  
may contain greater or fewer amino acid residues. A typical signal  
peptide consists of three regions: a basic N-terminal region, a central  
hydrophobic region, and a more polar C-terminal region. The central  
hydrophobic region contains 4 to 12 hydrophobic residues that anchor  
20 the signal peptide across the membrane lipid bilayer during transport  
of the nascent polypeptide. Following initiation, the signal peptide  
is usually cleaved within the lumen of the endoplasmic reticulum by  
cellular enzymes known as signal peptidases. Potential cleavage sites  
of the signal peptide generally follow the "(-3, -1) rule". Thus a  
25 typical signal peptide has small, neutral amino acid residues in  
positions -1 and -3 and lacks proline residues in this region. The  
signal peptidase will cleave such a signal peptide between the -1 and  
+1 amino acids. Thus, the portion of the DNA encoding the signal  
sequence may be cleaved from the amino-terminus of the immunofusin  
30 protein during secretion. This results in the secretion of a  
immunofusin protein consisting of the Fc region and the target protein.  
A detailed discussion of signal peptide sequences is provided by von  
Heijne (1986) Nucleic Acids Res., 14:4683 (incorporated herein by  
reference). As would be apparent to one of skill in the art, the  
35 suitability of a particular signal sequence for use in the secretion  
cassette may require some routine experimentation. Such

- 12 -

experimentation will include determining the ability of the signal sequence to direct the secretion of an immunofusin and also a determination of the optimal configuration, genomic or cDNA, of the sequence to be used in order to achieve efficient secretion of immunofusins. Additionally, one skilled in the art is capable of creating a synthetic signal peptide following the rules presented by von Heijne, referenced above, and testing for the efficacy of such a synthetic signal sequence by routine experimentation. A signal sequence is also referred to as a "signal peptide", "leader sequence" or "leader peptides" and each of these terms having meanings synonymous to signal sequence may be used herein.

The Fc region of an immunoglobulin is the amino acid sequence for the carboxyl-terminal portion of an immunoglobulin heavy chain constant region. The Fc regions are particularly important in determining the biological functions of the immunoglobulin and these biological functions are termed effector functions. As known, the heavy chains of the immunoglobulin subclasses comprise four or five domains: IgM and IgE have five heavy chain domains, and IgA, IgD and IgG have four heavy chain domains. The Fc region of IgA, IgD and IgG is a dimer of the hinge-CH2-CH3 domains, and in IgM and IgE it is a dimer of the hinge-CH2-CH3-CH4 domains. Further the CH3 domain of IgM and IgE is structurally equivalent to the CH2 domain of IgG, and the CH4 domain of IgM and IgE is the homolog of the CH3 domain of IgG (see, W.E.Paul, ed., 1993, Fundamental Immunology, Raven Press, New York, New York, which publication is incorporated herein by reference). Any of the known Fc regions would be useful as the Fc region of the secretion cassette. However, it is important that the binding sites for certain proteins be deleted from the Fc region during the construction of the secretion cassette. For example, since coexpression with the light chain is unnecessary, the binding site for the heavy chain binding protein, Bip (Hendershot et al. (1987) Immunol. Today 8:111-114), should be deleted from the CH2 domain of the Fc region of IgE, such that this site does not interfere with the efficient secretion of the immunofusin. Likewise, the cysteine residues present in the Fc regions which are responsible for binding to the light chain of the

- 13 -

immunoglobulin should be deleted or substituted with another amino acid, such that these cysteine residues do not interfere with the proper folding of the Fc region when it is produced as an immunofusin. In the same manner, transmembrane domain sequences, such as those present in IgM, should be deleted such that these sequences do not result in misdirecting the immunofusin to the membrane as a transmembrane protein.

Upon expression and production of the Fc region as a portion of the secretion cassette, it may retain some of the biological properties, termed "effector functions", which are native to the particular immunoglobulin class from which the Fc region is obtained. Useful effector functions include, for example, complement fixation, Fc receptor binding, binding to cell membranes, and placental transfer. In some cases, it may be advantageous to modify or remove one or more of these effector functions, such as Fc receptor binding or complement fixation, using site directed mutagenesis or other well known molecular biology techniques. For example, Duncan et al. (Nature, 1988, 332:738) have mapped the amino acids responsible for the several of the immunoglobulin gamma effector functions activities, see also, Duncan et al., 1988, 332:563; Yasmeen et al., Immunol., 1976, 116:518; Tao et al., J. Immunol., 1989, 143:2595. Each of these publications is incorporated herein by reference. The amino acids or peptide segments responsible for these functions can be deleted thus removing that portion of the Fc region, or substituted with sequences which would not confer the function using well known molecular biology techniques.

The currently preferred class of immunoglobulin from which the Fc region is derived is immunoglobulin gamma-1, because it has been well characterized and is efficiently secreted from most cell types. The Fc region of the other subclasses of immunoglobulin gamma (gamma-2, gamma-3 and gamma-4) would function equally well in the secretion cassette. The Fc region of immunoglobulin gamma-1 is preferably used in the secretion cassette includes at least part of the hinge region, CH2 region, and CH3 region. In addition, the Fc region of immunoglobulin gamma-1 can be a CH2-deleted-Fc, which includes a part of a hinge

- 14 -

region and a CH3 region wherein the CH2 region has been deleted. A CH2-deleted-Fc has been described by Gillies et al., 1990, Hum. Antibod. Hybridomas, 1:47, which publication is incorporated herein by reference.

5

As is apparent from the above discussion of Fc regions, the Fc regions from the other classes of immunoglobulins, IgA, IgD, IgE, and IgM, would also be useful as the Fc region of the secretion cassette. Further, deletion constructs of these Fc regions, in which one or more of the constant domains are deleted would also be useful. One of ordinary skill in the art could prepare such deletion constructs using well known molecular biology techniques.

The identity of the target protein produced in accordance with the invention is essentially unlimited. Indeed, an important feature of the invention is that it provides a generalized DNA construct, and procedure which can be adapted to facilitate recombinant production of any desired target protein. For instance, the application of the invention to the expression of the regulatory proteins, such as transcription factors which are normally localized to the nucleus, allows for the efficient secretion of such normally non-secreted proteins. In addition, regulatory proteins are in general difficult to express and the purification procedures are generally cumbersome (see, for example, Meisterernst et al. (1991) Cell 66:981). Therefore, it is especially desirable that such proteins be exported into the culture medium. Additionally, the invention can be used to enhance the production and secretion of proteins which are normally secreted at low levels. If a desired target protein includes sequences encoding a secretion signal or a transmembrane signal, these sequences can be removed from the target protein such that the secretion cassette directs the secretion of the fusion protein.

The optional proteolytic cleavage site may be any amino acid sequence which is recognized by specific cleavage agents. The specificity of cleavage agents is determined by the identity of the sequence of amino acids at or near the peptide bond which is to be



- 15 -

hydrolyzed. A given cleavage agent may recognize the bond between two specific amino acids or may recognize a bond following one or a specific sequence of amino acids. The specificity of many cleavage agents is known. Table 1 set forth below lists various known cleavage agents and their primary (and in some cases secondary) sites of action.

TABLE 1

<u>Cleavage Agent</u>	<u>Major Site of Action</u>	<u>Other Sites of Action</u>
10 Trypsin	Arg, Lys	
Chymotrypsin	Trp, Phe, Tyr	Leu, Met, His
Elastase	Neutral Aliphatic Residues	
Pepsin	Phe, Leu, Trp	Ala, Gly, Glu
15 Papain	Arg, Lys, Gly	Wide specificity
Subtilisin	Aromatic and Aliphatic residues	Various
Thermolysin	Amino-linked bonds of Aliphatic Residues	Ala, Phe
20 <i>S. aureus</i> protease	Glu	Asp
Endoproteinase Arg C (Submaxillaris protease)	Arg	
Clostripain	Arg	
25 Thrombin	Arg	
Collagenase	X-Gly-Pro	X-Ala-Pro X-Gly-Thr
Lysobacter enzymogenes	Lys	
30 (endoproteinase Lys-C)		
Mysobacter Al-1 Protease	Lys	
Armillaria mellea	Lys	
Flavobacterium	Pro	
35 meringosepticum		
Factor Xa	Ile-Glu-Gly-Arg	

- 16 -

	CNBr	Met
	BNPS-skatole	Trp
	N-bromosuccinimide	Trp
	O-iodosobenzoic	Trp
5	acid	
	HBr/DMSO	Trp
	NTCB	Cys
	Sodium metal in	
	liquid ammonia	Pro
10	Hydroxylamine	Asn-Gly
	Dilute acid	Asp-Pro

Other cleavage agents are known. Those preferred for use in the invention are enzymes with a primary site of action which cleave at the C-terminal side of the cleavage site residue.

The cleavage site in the fused protein generally can comprise any one or sequence of amino acids which can be cleaved by a cleavage agent specific for the site in an appropriate environment. Specificity of cleavage can be increased, and likelihood of undesired cleavage within the target protein or elsewhere in the fused polypeptide can be decreased, by selecting the cleavage agent having a site of action which is absent from the target polypeptide. The fused polypeptide is preferably cleaved under conditions in which it has assumed its native conformation. This has the effect of masking the presence of potential cleavage sites in the target polypeptide.

The invention is illustrated further by the following non-limiting examples.

30

#### Example 1. Construction of a Secretion Cassette

The construction of an exemplary secretion cassette is described below. As would be appreciated by those of ordinary skill in the art, the signal sequence and the Fc region of an immunoglobulin could be other sequences than those described.

- 17 -

The signal sequence of an immunoglobulin light chain of the 14.18 antibody was selected for use as the signal sequence of the secretion cassette. The sequence of the 14.18 antibody light chain is provided in Gillies et al., 1989, Jour. Immunol. Meth., 125:191-202 and is incorporated herein by reference. The signal sequence was modified for ease of cloning as an XbaI-AflIII fragment of the DNA. As would be apparent to those of skill in the art, the DNA encoding a human signal sequence could also be used. Specifically, an XbaI site was introduced 5' of the translation initiation codon and the consensus sequence for optimal ribosome binding (Kozak, 1984, Nature 308:241, incorporated herein by reference). An AflIII site was introduced into the 3' end of the signal sequence by mutagenizing the DNA coding for the penultimate amino acid residue of the signal peptide from a serine to a leucine, thus the sequence ATC was mutagenized to TTA using site directed mutagenesis.

The Fc region of an immunoglobulin was selected to be the human Fc $\gamma$ 1 genomic DNA, including the genomic configuration of the hinge, CH2 and CH3 domains. The genomic sequence of human Fc $\gamma$ 1 is provided in Huck et al., (1986) Nucleic Acids Res. 14:1779 and is incorporated herein by reference. As would be apparent to one of ordinary skill in the art, a CH2-deleted-Fc may also be used as the Fc region of the secretion cassette (see, Gillies et al., 1990, Hum. Antibod. Hybridomas, 1:47), in which case the CH2 domain would be deleted from the Fc region using established molecular biology techniques during the construction of the secretion cassette. The genomic DNA of Fc $\gamma$ 1 was modified for ease of cloning as an AflIII-XmaI fragment. The 5' end of the human Fc genomic DNA was mutagenized to an AflIII site by performing a Polymerase Chain Reaction (PCR) using a 5' sense primer with the following sequence (Sequence ID No. 1):

GAGAATTCTTAAGCGAGCCCAAATCTTCTGACAAACTCAC

This primer introduced an AflIII site (underlined) and a cysteine to serine mutation (TGT to TCT, bold). The cysteine being mutated is the one that is normally involved in disulphide bonding with the light chain and thus does not affect the effector functions of the Fc region.

- 18 -

The deletion of this cysteine may serve to enhance the production of the Fc $\gamma$ 1 region as the efficient production of this modified Fc $\gamma$ 1 region will not require the coexpression of the immunoglobulin light chain. This cysteine was also removed such that it does not interfere with the proper folding of the Fc $\gamma$ 1 region or the fused target protein. The 3' end of the Fc $\gamma$ 1 genomic DNA encodes for two XmaI restriction sites. They are located at 10 and 280 bp upstream of the translation stop codon in the CH3 domain. The distal XmaI site was destroyed by introducing a silent mutation, using site directed mutagenesis, (TCC to TCA, where the CC were the first two bases of the XmaI site) so that the XmaI site 10 bp upstream of the stop codon became unique.

The XbaI-AflIII restriction fragment encoding the light chain signal peptide was then ligated to the AflIII-XmaI restriction fragment encoding the Fc region. The resultant XbaI-XmaI restriction fragment therefore encodes the secretion cassette, and the gene encoding the target protein of interest can be ligated to the 3' end of the secretion cassette via the XmaI site.

In general, the DNA encoding the target protein can be ligated to the unique XmaI site through the use of a linker-adaptor, such a linker-adaptor may also include restriction endonuclease sites in addition to an XmaI site. The use of a linker-adaptor has the additional feature in that it can encode a proteolytic cleavage site for subsequent use in cleaving the target protein from the secretion cassette after production and secretion of the fusion protein. For example, the linker-adaptor can encode a lysine residue at the junction of the fusion protein, which provides the option of cleaving the target protein from the Fc domain by proteolytic enzymes such as trypsin or plasmin. Similarly, the linker adaptor can include a DNA encoding the cleavage site of enterokinase K (Asp-Asp-Asp-Asp-Lys) in order to provide for the specific cleavage of the secreted fusion protein by enterokinase K.

Example 2. Construction of an Immunofusin

- 19 -

The construction of an exemplary immunofusin, including a secretion cassette and a target protein is described below. As would be apparent to those of ordinary skill in the art, other target proteins can be fused to a secretion cassette using the same or other molecular cloning techniques.

The target protein for the exemplary immunofusin was chosen to be CD26, which is a type II membrane protein having its active site within the carboxyl-terminal region of the protein which is the extracellular domain. During the construction of a CD26 immunofusin, the cytoplasmic and transmembrane domains of CD26 were deleted so that they would not interfere with the secretion of the immunofusin by the secretion cassette. The 5' end of the cDNA encoding the extracellular domain was modified for ease of cloning to include a XmaI site, which was introduced via a linker-adaptor. The 3' end of the CD26 cDNA was also modified for ease of cloning to include a XhoI site, which could be introduced downstream of the translation stop codon either by PCR or by linker-adaptor ligation.

Various linker-adaptors can be used depending upon the desire for introduction of a proteolytic cleavage site between the DNA encoding for the Fc region and the CD26 cDNA. For example, one linker-adaptor which can be used for CD26 is:

5' CCG GGT (AAA) GGC ACA GAT GAT GCT ACA G  
3' CA (TTT) TTG TGT CTA CTA CGA TGT C

as provided in Sequence ID Nos. 2 and 3. The first three codons in the top strand encode the last three amino acid residues of the CH3 domain, and starting with the codon GGC is the gene sequence of the extracellular domain of CD26. This linker-adaptor had the cohesive end of an XmaI site at its 5' end and the blunt end of a PvuII site at its 3' end, the blunt ended PvuII site being a convenient site for reconstruction with the rest of the CD26 cDNA. The lysine codon (AAA, in parenthesis) in the linker-adaptor is but one of many optional amino acid sequences which are useful to provide for a proteolytic cleavage site by cleavage agents. For example, this lysine residue can be cleaved by enzymes such as trypsin or plasmin.

- 20 -

Alternatively, for more specific proteolytic cleavage by enterokinase K, the gene sequence encoding the enterokinase K cleavage site can be introduced via the following linker-adaptor:

5           5' CCG GGT TCA GGG GAT **GAC GAT GAC GAT A**  
          3'       CA AGT CCC CTA CTG CTA CTG CTA TTC GA

as provided in Seq. ID Nos. 4 and 5. The nucleotides in bold encode the amino acid residues (Asp)4-Lys, which is the recognition site of enterokinase K. The linker-adaptor ends with a HindIII site, to which  
10 the CD26 gene or other target protein gene sequences can be joined.

### Example 3. Host Cells and Transfection

The preferred host cell lines include the mouse myeloma (or  
15 hybridoma) NS/0 and Sp2/0 Ag14 cells. The myeloma cells were transfected by protoplast fusion and selected in Dulbecco's modified Eagle's medium (Gibco) containing 10% fetal bovine serum and 100 nM methotrexate, as described by Gillies et al., 1989, BioTechnology, 7:799, which publication is incorporated herein by reference.  
20 Transfectants secreting the immunofusins were identified by anti-Fc ELISA, as described by Gillies et al. (1989) J. Immunol. Methods 125:191, which publication is incorporated herein by reference. The highest producers were adapted to media containing 1 µM MTX and subcloned by limiting dilutions. For the production of immunofusins,  
25 the cells were grown in Hybridoma Serum-Free Media (HSFM, Gibco) containing 1% fetal bovine serum and 1 µM MTX.

The other preferred recipient cell line is the human kidney 293 cells, which is useful for both transient and stable expression. Other  
30 cells, such as the HeLa and the Chinese hamster ovary (CHO) cells, also worked in our system. The preferred method of transfection for these adherent cells is by coprecipitation of plasmid DNA with calcium phosphate, and other methods include lipofection and electroporation. For a description of these methods and other useful transfection  
35 methods see, Sambrook et al. (1989) Molecular Cloning--A Laboratory Manual, Cold Spring Harbor, NY, incorporated herein by reference.

- 21 -

Example 4. Characterization and Purification of  
Immunofusins

5 For routine characterization by gel electrophoresis, immunofusins in the conditioned media were first captured on Protein A Sepharose (Repligen, Cambridge, MA) and then eluted by boiling in protein sample buffer with or without 2-mercaptoethanol. After electrophoresis on an SDS-gel, the protein bands were visualized by Coomassie staining. For  
10 example, the IL2 immunofusin, see example 5, gave a band having the molecular weight of 45 kD under reducing conditions and a band having the molecular weight of 90 kD under non-reducing conditions, showing that the IL2 immunofusin was produced as a dimer, presumably through disulphide bonding in the hinge domain of the Fc region.

15 For purification, the cell culture media was collected and then the immunofusins were bound on Protein A Sepharose. The immunofusins were subsequently eluted from the Protein A in a sodium citrate buffer (100 mM, pH 4). The eluate was then immediately neutralized with 0.1 volume  
20 of 1 M Tris-hydrochloride, pH 8. In the case of CD26 immunofusin, it was shown that such an elution procedure resulted in greater than 80% recovery of the CD26 immunofusin with no loss of enzyme activity.

Example 5. Expression of IL2 Immunofusin

25 The cDNA of mature IL2 protein was modified for ease of cloning to have a 5' XmaI restriction endonuclease site and a 3' XhoI restriction endonuclease site using well known molecular techniques, such as those which were as described in example 2. The sequence of the mature IL2  
30 cDNA is provided in Taniguchi et al., 1983, Nature, 302:305 and is incorporated herein by reference. The cDNA of the mature IL2 protein was constructed using recombinant techniques as a synthetic gene in order to optimize codon usage and to introduce desirable restriction endonuclease cleavage sites. The synthetic gene was created using  
35 conventional DNA manipulation techniques. Once the synthetic IL2 cDNA was constructed, the 5' XmaI site of the IL2 cDNA was ligated to the 3'

- 22 -

XmaI site of the secretion cassette, described in Example 1. The IL2 immunofusin, was then cloned into the expression vector pdC. The IL2 immunofusin expression vector was transfected into NS/0 and Sp2/0 as host cells by protoplast fusion, as is described by Gillies et al., 1989, Biotechnology, 7:799.

Two to three weeks after transfection, MTX-resistant NS/0 and Sp2/0 clones appeared. The initial clones were screened by anti-Fc ELISA. The IL2 immunofusin protein was collected from the media. An appropriate assay for the biological activity of IL2 was the standard T-cell proliferation assays according to Gillies et al. (Proc. Natl. Acad. Sci. (1992) 89:1428), which is incorporated herein by reference. The spent culture of the best clone contained about 100 µg/ml of IL2 immunofusin. The host cell clones which efficiently produced and secreted the IL2 immunofusin protein were subcloned in media containing 100 nM MTX, and the best subclone produced about 200 µg/ml of protein in spent culture. When MTX was left out of the media in the subcloning, the best subclone thus isolated produced about 180 µg/ml in spent culture. Thus, the construction of an IL2 immunofusin unexpectedly provided for the production of IL2 at a level which is about 80 times that which can be achieved by the expression of IL2 alone using the pdEMP vector (unpublished data), and many times of that of the IL2 that was expressed in mammalian cells (Conradt et al., 1989, J. Biol. Chem., 264:17368) and in yeast (Ernst et al., 1989, Biotechnology, 7:716). As mentioned in example 4, IL-2 immunofusin was produced as a homo-dimer of molecular weight of 90 kD, presumably through disulphide bonding in the hinge domain of the 45 kD monomers.

#### Example 6. Expression of CD26 immunofusin

30

The construction of CD26 as an immunofusin was undertaken to demonstrate that the invention is applicable to the expression of membrane anchored proteins such as type II membrane proteins. A type II membrane protein displays the carboxyl-terminal domain on the extracellular surface, and most often includes its active region within this carboxyl-terminal domain. The joining of a fusion polypeptide to



- 23 -

the carboxyl-terminal region of such a protein may interfere with the proper folding of the active site, and thus reduce or prevent the production of active protein.

5 CD26 is a type II membrane protein comprising 766 amino acid residues. The biological function of CD26 is as a T cell activation antigen and the putative coreceptor for entry of HIV in CD4+ cells (Callebaut et al. (1993) Science 262:2045). The CD26 protein is anchored to the lipid bilayer of the plasma membrane through a  
10 hydrophobic domain between residues 7 and 28 at the N-terminus. Amino acids 1 to 6 form a short cytoplasmic tail. The rest of the protein, between residues 29 and 766, is extracellular and includes several potential N-glycosylation sites and the active site of the enzyme (Tanaka et al. (1992) J. Immunol. 149:481). The 728 carboxyl-terminal  
15 residues in CD26 protrude from the membrane surface and the C-terminus is free. A soluble CD26 expressed as an immunoadhesin, will have a conformation different from that of the native CD26, because the carboxyl-terminus in an immunoadhesin CD26 protein is not free but connected to antibody sequence. On the other hand, if we engineer an  
20 immunofusin in which the antibody sequence is amino-terminal to the target protein, such as CD26, the native conformation of CD26 will be preserved, i.e. the C-terminus is free, and the antibody sequence, herein an Fc region, takes the place of the membrane to which CD26 is normally anchored. The enzymatic and biological activities of such a  
25 soluble CD26 immunofusin will not be compromised. In addition, CD26 is a protease and its expression may be deleterious to the host cell. Thus by efficiently exporting the CD26 protease outside of the host cell in the form of an immunofusin, a higher level of expression can be achieved.

30

A 2.3 kb cDNA fragment encoding the extracellular domain of CD26 was used to construct the CD26 immunofusin expression vector. The DNA sequence of CD26 is provided in Tanaka et al., 1992, J. Immunol., 149:481 and is incorporated herein by reference. CD26 was fused 3' of  
35 the secretion cassette as described above in example 2, and then the secretion cassette and CD26 target protein were cloned into the

- 24 -

expression vector pdc using the XbaI restriction endonuclease site 5' of the light chain signal sequence and the XhoI restriction endonuclease site 3' of the CD26 protein as described in example 2 above. The resultant CD26 immunofusin expression vector was

5 transfected into a host cell as described in Example 3 above. MTX-resistant clones from transfected NS/0 and Sp2/0 cells were screened by anti-Fc ELISA and DPPIV activity assay. CD26 is also known as DPPIV, which is an exopeptidase that cleaves after amino-terminal X-P (X can be any amino acid residue, and P is proline). DPPIV enzyme activity of

10 the CD26 immunofusin was assayed according to Tanaka et al., Proc. Natl. Acad. Sci., 1993, 90:4586, incorporated herein by reference, using glycylproline p-nitroanilide tosylate (Gly-Pro-pNA) as a substrate. The best NS/0 clone produced about 3.5 µg/ml of CD26 immunofusin. The DPPIV moiety of the protein product was determined to

15 be fully active, having  $K_M$  and  $k_{cat}$  values similar to those of the native CD26. Furthermore the enzymatic activity of CD26 immunofusin was inhibited by known peptide inhibitors in a dose-dependent manner. The peptide inhibitors tested included the tripeptides IPI and VPL and APL, each of which inhibited the CD26 enzyme activity greater than 30%

20 at 0.15 mM, greater than 70% at 1 mM and greater than 90% at 4 mM. As a control known non-inhibitor peptides were also tested for their effect upon CD26 enzyme activity and the known non-inhibitors, GGG and GPHyP (wherein HyP is hydroxproline), were found to have no effect on the CD26 activity when incubated with the CD26 immunofusin at

25 concentrations ranging between 0.01 mM and 11 mM.

#### Example 7. Expression of Tat immunofusin

The invention was also applied to the expression of regulatory

30 proteins which are normally localized to the nucleus. Because regulatory proteins are in general difficult to express and purify, it is especially desirable to devise a method by which such proteins can be efficiently secreted from a host cell. Immunofusin constructs of Tat and Rev (described in example 8), which are two proteins encoded by

35 the human immunodeficiency virus (HIV) that regulate expression of viral proteins in the cell nucleus, were made in order to determine the

- 25 -

efficiency with which these proteins can be expressed and collected. We obtained high level expression and secretion of the Tat and Rev immunofusins, and readily purified the immunofusins in a single step.

A 260 base-pair cDNA fragment encoding Tat was cloned into the XmaI and XhoI sites of the pDC expression vector by modification of the 5' and 3' ends of the Tat protein using recombinant DNA techniques as described above. The sequence of the cDNA encoding the Tat protein is provided in Ratner et al., 1985, Nature, 313:277, and is incorporated herein by reference. Specifically, the sequence at the 5' end was modified to, Seq. ID No. 6, C CCG GGT CGC ATG GAG . . . . , where the underlined sequence is the XmaI site and the ATG in bold is the translation start codon of the Tat gene. At the 3' end, an XhoI site was introduced immediately downstream of the translation stop codon by standard PCR techniques. The Tat immunofusin expression vector was then transfected into a host cell, as described above, and the host cells were analyzed for production of Tat immunofusin protein. High level expression was obtained in transiently transfected 293 cells and stably transfected NS/O cells. Stable NS/O clones produced about 3 µg/ml of a 48 kD protein, analyzed on a SDS-gel under reducing conditions. This protein was confirmed to be Tat immunofusin by an anti-Tat antibody (Cat. #7001, American BioTechnologies, Cambridge, MA).

The Tat immunofusin was shown to be active by the following transient expression experiment in 293 cells, the results of which are presented below in Table 2. The expression vector for Tat immunofusin was cotransfected with a separate vector containing LTR-TAR-Kappa, where LTR-TAR is the long terminal repeat DNA sequence of HIV that is transactivated by the Tat protein, and Kappa is the gene sequence encoding the Kappa light chain of immunoglobulin. To measure expression levels of Fc-Tat (Tat immunofusin) and Kappa, the supernatants were assayed by anti-Fc and anti-Kappa ELISA respectively. In Table 2, pDC-Fc-Tat represents the pDC expression vector for Tat immunofusin; LTR-TAR-Kappa represents the expression vector for Kappa light chain, in which the LTR-TAR regulatory region can be transactivated by Tat; and pCEP-Tat is an expression vector for Tat,

- 26 -

whose transcription is under the control of the human cytomegalovirus enhancer and promoter. pCEP-Tat was used as a positive control to monitor the transactivation of the LTR-TAR-Kappa by Tat protein. As a negative control LTR-TAR-Kappa was transfected alone to demonstrate that it is not transactivated in the absence of Tat protein or Tat immunofusin. As shown in Table 2, high level expression of the Tat immunofusin was observed in transfection 1; high level expression of both Tat and Kappa light chain were observed in the cotransfection experiment, transfection 2. Transactivation of Kappa by Tat was seen in the positive control, transfection 3, as expected. Little or no expression of Kappa was seen in the negative control, transfection 4, also as expected. Therefore, the Kappa light chain is expressed only through transactivation of the LTR-TAR region by a functional Tat protein, and the Tat immunofusin provides a functional Tat protein which is readily secreted from the host cell. This result also demonstrates that the secretion cassette is able to direct the secretion of a protein which is normally transported to the nucleus of the host cell.

Table 2

DNA used in transfection	ELISA (ng/ml)	
	Fc	Kappa
1.) pdC-Fc-Tat	>3000	0
2.) pdC-Fc-Tat, LTR-TAR-Kappa	1600	160
3.) PCEP-Tat, LTR-TAR-Kappa	0	277
4.) LTR-TAR-Kappa	0	3

#### Example 8. Expression of Rev Immunofusin

A 350 base-pair cDNA fragment encoding Rev was modified to include a 5' XmaI site and a 3' XhoI site and then ligated 3' of the described secretion cassette in the pdC expression vector. The sequence of the cDNA encoding the Rev protein is provided in Ratner et al., 1985, Nature, 313:277, and is incorporated herein by reference. Specifically, the 5' end of the cDNA was modified to C CCG GGT CGC **ATG** GCA . . . . (Seq. ID No. 7), where the underlined sequence is the XmaI site and the ATG in bold is the translation start codon of the Rev

- 27 -

gene. At the 3' end, an XhoI site was introduced immediately downstream of the translation stop codon by standard PCR techniques. High level expression was obtained in transiently transfected 293 cells and stably transfected NS/O cells. Stable NS/O clones produced about 3  $\mu\text{g}/10^6$  cells/day of the Rev immunofusin, which has a molecular weight of about 50 kD when analyzed on a SDS-gel under reducing conditions.

Example 9. Site-specific Proteolytic Cleavage of an Immunofusin

An exemplary cleavage of an immunofusin is described below, as would be apparent to one of ordinary skill in the art, each of the above described immunofusins could be cleaved from their respective secretion cassettes using the same method or an analogous method.

A CD26 immunofusin having a lysine residue ("Fc(Lys)-CD26 immunofusin"), introduced by linker adaptor during construction of the immunofusin between the Fc region and the CD26 target protein sequence was cleaved using trypsin. To cleave the Fc(Lys)-CD26 immunofusin, the immunofusin was bound on Protein A Sepharose and cleaved at the desired lysine position by trypsin to release CD26 as follows: Fc(Lys)-CD26 immunofusin bound on Protein A Sepharose was incubated with a 1% trypsin solution at 37°C for 2 hr. Trypsin inhibitor (Sigma) was then added to stop any further digestion. The supernatant was then removed and analyzed on an SDS-gel under reducing conditions. After Coomassie staining, a band having a molecular weight of 110 kD, which corresponds to the size of CD26 without the secretion cassette, was obtained. As a control, CD26 immunofusin, without the lysine residue at the junction of the fusion between the Fc domain and the CD26 target protein ("Fc-CD26 immunofusin"), was bound on Protein A Sepharose and similarly treated. The CD26 was found to not be released from the secretion cassette of the Fc-CD26 immunofusin, as was expected, and this also confirmed the specific cleavage of the immunofusin at the amino acid lysine which was inserted between the CH3 domain of the Fc region and the target CD26 protein. As a further control, an identical aliquot of Fc-CD26 immunofusin which was bound to Protein A Sepharose was boiled

- 28 -

in the protein sample buffer and SDS-gel analysis of the supernatant showed a 140 kD band corresponding to the full length CD26 immunofusin protein monomer.

- 5       The results from the gel electrophoresis experiment were confirmed by DPPIV activity assays of the tryptic digests. Quantitative recovery of the DPPIV enzymatic activity was obtained in the supernatant when the Fc(Lys)-CD26 immunofusin bound to Protein A Sepharose was treated with trypsin. In the parallel experiment with Fc-CD26 immunofusin,  
10       there was no DPPIV activity in the supernatant, because the CD26 protein was not released from the Protein A Sepharose.

Example 10. Expression of OSF-2 immunofusin

- 15       OSF-2 is a 80-kD secretory protein that is involved in the ossification process. The sequence the DNA encoding OSF-2 is provided in Takeshita et al., 1993, Biochem. J. 294:271, and is incorporated herein by reference. The cDNA encoding the OSF-2 protein with its signal peptide was cloned into the expression vector pdC. NS/O cells  
20       were used for stable transfection and 293 cells were used for transient expression; but in neither case was the OSF-2 protein detected.

- The OSF-2 cDNA was then adapted to be expressed as an immunofusin. At the 3' end, the XbaI site at the translation stop codon was  
25       converted to an XhoI site by linker ligation. At the 5' end the following linker-adaptor was used:

5'    CCG GGT AAA AAC AAT CAT TAT GAC AA

3'           CA TTT TTG TTA GTA ATA CTG TTC TAG

- as provided in Seq. ID Nos. 8 and 9. The nucleotides in bold encode  
30       the N-terminus of the mature OSF-2 protein, ending with BglII cohesive ends. These BglII cohesive ends were ligated to the BglII-XhoI fragment of the OSF-2 cDNA. The XmaI cohesive ends at the 5' end of the linker-adaptor (underlined) were ligated to the unique XmaI site in the immunofusin expression vector.

35

- 29 -

High level expression was obtained in transiently transfected 293 cells and stably transfected NS/0 cells. Stable NS/0 clones produced about 5 to 7 µg/ml of a 110 kD protein, when analyzed on a SDS-gel under reducing conditions. This protein was confirmed to be the OSF-2 immunofusin by Western blotting with an anti-OSF-2 antibody.

It was also found that the expression of OSF-2 as an immunofusin in a mammalian system was superior to the expression of OSF-2 in the thioredoxin gene fusion expression system in *E. coli* (LaVallie et al., 1993, *Biotechnology*, 11:187). The thioredoxin gene fusion system was designed to circumvent the formation of inclusion bodies because fusion to thioredoxin increases the solubility of many heterologous proteins produced in the *E. coli* cytoplasm. To test this system for the expression of OSF-2, the cDNA encoding the mature OSF-2 was inserted into the SmaI site of the pTrxFus vector (Invitrogen, San Diego, CA), thus creating a thioredoxin OSF-2 fusion protein. The supplier's protocol for the expression of the fusion proteins was followed. The thioredoxin OSF-2 fusion protein was expressed, and, as a control, the thioredoxin protein was expressed alone without a fusion partner. The results showed that although thioredoxin alone could be produced as a soluble protein at a high level, the thioredoxin OSF-2 fusion protein was present only in the insoluble fraction. Therefore, in addition to the lack of post-translational modification in bacterial expression, a relatively complex mammalian protein such as OSF-2 was not synthesized as a soluble protein when fused to thioredoxin.

#### Example 11. Expression of $\beta$ IG-H3 immunofusin

$\beta$ IG-H3, a gene product which is induced by transforming growth factor- $\beta$ , is a 68-kD secretory protein that shares sequence homology with OSF-2. The sequence of cDNA encoding  $\beta$ IG-H3 is provided in Skonier et al. (1992) *DNA and Cell Biology*, 11:511, and is incorporated herein by reference. The cDNA encoding the native  $\beta$ IG-H3 was cloned into the expression vector pdC; but attempts to obtain stable transfectants producing  $\beta$ IG-H3 were unsuccessful.

- 30 -

The  $\beta$ IG-H3 cDNA was then adapted to be expressed as an immunofusin. At the 3' end, the BsmI site downstream of the translation stop codon was converted to an XhoI site by linker ligation. At the 5' end, the following linker-adaptor was used:

5 5'           CCG GGT AAA GCC CTG GGC C  
3'           CA TTT CGC GAC

(Seq ID. Nos. 10 and 11). The nucleotides in bold encode the N-terminus of the mature  $\beta$ IG-H3 protein. The linker-adaptor had XmaI cohesive ends for ligating to the expression vector as described in the above examples, and ApaI cohesive ends for ligating to the ApaI site at the 5' end of the cDNA sequence encoding the mature  $\beta$ IG-H3.

High level expression was obtained in transiently transfected 293 cells and stably transfected NS/0 cells. Stable NS/0 clones produced about 3.5  $\mu$ g/ $10^6$  cells/day of a 100 kD protein when analyzed on a SDS-gel under reducing conditions. This protein was confirmed to be the  $\beta$ IG-H3 immunofusin by Western blotting with anti- $\beta$ IG-H3 antibody.

20 Example 12. Expression of the soluble form of IgE receptor as an immunofusin

The high affinity IgE receptor alpha subunit (IgE-R), the DNA sequence of which can be found in Kochan et al. (1988) Nucleic Acids Res. 16: 3584 and is incorporated herein by reference, was constructed as an immunofusin as follows: An XmaI site was introduced to the 5' end of the cDNA encoding the mature IgE-R so that the sequence at the junction of the fusion was C CCG GGT GTC CCT CAG --- (Seq. ID No. 12), where the XmaI site is underlined and the three codons in bold are the first three amino acid residues of the mature IgE-R. At the 3' end of the IgE-R, the cDNA encoding the transmembrane domain and the rest of the C-terminus was deleted and a translation stop codon was placed after the last codon of the extracellular domain. The sequence of the IgE-R immunofusin at the 3' end was thus TAC TGG CTA TAA CTC GAG (Seq. ID No. 13), where the three codons in bold were the last three amino acid residues of the extracellular domain of the IgE-R, and they were followed by a stop codon and an XhoI site (underlined).



- 31 -

The pdC expression vector containing the IgE-R immunofusin was transfected into 293 cells and NS/0 cells. High levels of expression (3 to 5 µg/ml) of the IgE-R immunofusin were detected in the cell culture media by anti(Fc) ELISA. SDS-gel analysis under reducing conditions showed a band of the expected size of 70 kD. The partially purified protein (on Protein A Sepharose) was shown to bind IgE in an IgE-R/IgE ELISA.

10 Example 13. Expression of Fcγ1

Fcγ1 was expressed by itself without a C-terminal target protein. This was achieved by ligating the following linker (having XmaI and XhoI cohesive ends)

15 5' CCG GGT AAA TAG C

3' CA TTT ATC GAG CT

(Seq. ID Nos. 14 and 15), to the XmaI and XhoI sites of the pdC to reconstruct the coding region of Fc. High levels of expression was detected by anti(Fc) ELISA in the cell culture media of the transiently transfected 293 cells (5 to 7 µg/ml) and stably transfected NS/0 clones (5 to 10 µg/ml). SDS-gel analysis under reducing conditions showed an Fc band of the expected size of 31 kD.

Example 14. Expression of PSMA immunofusin

25

PSMA, prostate specific membrane antigen, is a type II membrane protein having a molecular weight of greater than 100 kD. PSMA is an integral membrane protein, and as such it is an attractive target for imaging and immunoconjugate delivery. To facilitate the expression of significant quantities of PSMA, we subcloned the extracellular domain of PSMA (the soluble form) and expressed this domain of PSMA as an immunofusin. A portion of the extracellular domain of PSMA, which is a soluble form of PSMA, can be produced as an immunofusin.

35

The cDNA encoding the full length PSMA was cloned from a human prostate carcinoma cell line LNCaP [Israeli et al. (1993) Cancer Res.,

- 32 -

53:227, which publication is incorporated herein by reference]. The portion of the PSMA cDNA corresponding to the extracellular domain was adapted to be expressed as an immunofusin by Polymerase Chain Reaction using the following primers:

5 N-terminal: 5' AAGCTT AAA TCC TCC AAT GAA GC

C-terminal: 5' CTCGAG TTA GGC TAC TTC ACT CAA AG

(Seq. ID Nos. 16 and 17). The two primers provide the HindIII and the XhoI sites (underlined) for cloning into the immunofusin expression vector. In the N-terminal primer, the HindIII site is followed by the coding sequence of the extracellular domain of PSMA (in bold) immediately after the transmembrane region. In the C-terminal primer, the XhoI site is followed by the anticodon of the STOP codon and the C-terminal coding sequence of PSMA (in bold). The amino acid sequence of the extracellular domain of PSMA is shown in Seq. ID No. 18.

15

High level expression was obtained in stably transfected 290 and Sp2/0 cells. The PSMA immunofusin secreted into the cell culture media was purified by Protein A Sepharose. Treatment of the immunofusin with the protease plasmin quantitatively converted the 130-kD Fc-PSMA into two products: the 100-kD PSMA extracellular domain and the 31-kD Fc. The Fc was then removed from the solution by adsorption onto Protein A Sepharose. The soluble PSMA was purified and used to immunize mice. It is expected that an antibody specific only to PSMA should facilitate diagnosis and therapy of prostate cancer.

25

#### Example 15. Expression of Murine Fc

The Fc region of murine  $\gamma 2a$  was prepared for expression as an immunofusin. Since the murine Fc region will not be immunogenic to mice, such an immunofusin containing the murine Fc followed by, for example, a human protein fusion partner can be used to immunize mice directly without prior cleavage to get rid of the Fc. The murine Fc was cloned into our immunofusin expression vector as described below, and was expressed at a high level under our expression conditions.

35

- 33 -

The murine Fc  $\gamma$ 2a domain, preceded by the signal peptide described above, was cloned into an expression vector, pdC, and was expressed without fusion to a target protein. Murine Fc  $\gamma$ 2a cDNA (Sikorav et al., 1980, Nucleic Acids Res., 8:3143-3155, which publication is incorporated herein by reference) was adapted for cloning into the expression vector by Polymerase Chain Reaction using the following primers:

N-terminal: 5' CTTAAGC GAG CCC AGA GGG CCC ACA  
C-terminal: 5' CTCGAGC TCA TTT ACC CGG AGT CCG

(Seq. ID Nos. 19 and 20). The N-terminal primer contains an AflIII site (underlined) for ligating to the AflIII site at the 3' end of the signal peptide, described above. The sequence following the AflIII site (in bold) encodes the amino acid residues in the hinge region of murine  $\gamma$ 2a gene. The C-terminal primer contains an XhoI site for cloning into the expression vector, followed by the anticodons of the translation STOP codon and the carboxyl end of murine  $\gamma$ 2a (in bold).

High level expression of the murine Fc $\gamma$ 2a region was demonstrated in 293 cells by SDS gel analysis followed by Western blotting with an anti-murine IgG antibody.

Example 16. Expression of gp120

The envelope protein gp120 of human immunodeficiency virus (HIV) is a glycoprotein having a molecular weight of 120kD, and is expressed on the surface of HIV particles and HIV infected cells. The protein gp120 is originally expressed in infected cells as a polyprotein, gp160, which is then cleaved by a cellular protein to gp120 and gp41. gp120 was prepared as an immunofusin and determined that the gp120 immunofusin was expressed at a very high level. Any desired portion of gp120 may also be prepared as immunofusin. The Fc moiety of the gp120 immunofusin could be cleaved off and gp120 was purified.

The complete nucleotide sequence of HIV has been published in Ratner et al. (1985) Nature, 313:277, and this publication is incorporated herein by reference. To prepare the gp120 immunofusin, a

- 34 -

translation STOP codon followed by an XhoI restriction site was introduced to the gp120-gp41 junction after amino acid Arg-518 of gp160 using standard molecular biology techniques, e.g., polymerase chain reaction. The existing NdeI restriction site present at nucleotide 5 5979, which is within the amino terminal portion of gp120, was converted to a HindIII restriction site through linker-adaptor ligation to generate and in-frame fusion. The resultant HindIII-XhoI fragment (1.36 kilobase pairs) encoding gp120 was then cloned into the immunofusin expression vector, pDC, as described above.

10

The gp120 immunofusin expression vector was expressed in stably transfected 293 cells according to the methods described above, and high level expression of the gp120 immunofusin was obtained. The gp120 immunofusin was functionally active, as determined by binding to CD4 in 15 an ELISA. The gp120 immunofusin was also determined to be quantitatively cleaved by enterokinase to release gp120 and the Fc region.

#### Other Embodiments

20

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The present embodiments are therefore to be considered in all respects as illustrative and non-restrictive, the scope of the invention being 25 indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

- 35 -  
SEQUENCE LISTING

- 5 (1) GENERAL INFORMATION:
- (i) APPLICANT:  
    (A) NAME: FUJI IMMUNOPHARMACEUTICALS CORP.  
    (B) STREET: 125 HARTWELL AVENUE  
    (C) CITY: LEXINGTON  
10     (D) STATE: MA  
    (E) COUNTRY: USA  
    (F) POSTAL CODE: 02173  
    (G) TELEPHONE: (617) 861-5300  
    (H) TELEFAX: (617) 861-5301  
15     (I) TELEX:
- (ii) TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF  
      PROTEINS AS IMMUNOFUSINS
- 20 (iii) NUMBER OF SEQUENCES: 20
- (iv) CORRESPONDENCE ADDRESS:  
    (A) ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
                  THIBEAULT  
25     (B) STREET: 125 HIGH STREET  
    (C) CITY: BOSTON  
    (D) STATE: MA  
    (E) COUNTRY: USA  
30     (F) ZIP: 02110
- (v) COMPUTER READABLE FORM:  
    (A) MEDIUM TYPE: Floppy disk  
    (B) COMPUTER: IBM PC compatible  
    (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
35     (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
    (A) APPLICATION NUMBER: US  
    (B) FILING DATE:  
40     (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
    (A) NAME: PITCHER, EDMUND R.  
    (B) REGISTRATION NUMBER: 27,829  
45     (C) REFERENCE/DOCKET NUMBER: FIP-001
- (ix) TELECOMMUNICATION INFORMATION:  
    (A) TELEPHONE: 617-248-7000  
50     (B) TELEFAX: 617-248-7100
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:  
55     (A) LENGTH: 41 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear
- 60 (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
65

- 36 -

GAGAATTCTT AAGCGAGCCC AAATCTTCTG ACAAACCTCA C  
41

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCGGGTAAAG GCACAGATGA TGCTACAG  
28

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGTAGCATC ATCTGTGTTT TTAC  
24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCGGGTTCAG GGGATGACGA TGACGATA  
28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- 37 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

5 AGCTTATCGT CATCGTCATC CCCTGAAC  
28

(2) INFORMATION FOR SEQ ID NO:6:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

20 CCCGGGTCGC ATGGAG  
16

(2) INFORMATION FOR SEQ ID NO:7:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCC GG GTCGC ATGGCA  
40 16

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
45 (A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

55 CCGGGTAAAA ACAATCATTG TGACAA  
26

(2) INFORMATION FOR SEQ ID NO:9:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
65 (D) TOPOLOGY: linear

- 38 -

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCTTGTCA TAATGATTGT TTTTAC  
26

10 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
15 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGGGTAAAG CCCTGGGCC  
25 19

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 base pairs  
30 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

40 CAGGGCTTTA C  
11

(2) INFORMATION FOR SEQ ID NO:12:

45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCCCGGGTGTC CCTCAG  
60 16

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
65 (B) TYPE: nucleic acid



- 39 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

10 TACTGGCTAT AACTCGAG  
18

(2) INFORMATION FOR SEQ ID NO:14:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGGGTAAAT AGC  
13

30 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 base pairs  
(B) TYPE: nucleic acid  
35 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

45 TCGAGCTATT TAC  
13

(2) INFORMATION FOR SEQ ID NO:16:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

60 AAGCTTAAAT CCTCCAATGA AGC  
23

(2) INFORMATION FOR SEQ ID NO:17:

65

- 40 -

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTCGAGTTAG GCTACTTCAC TCAAAG  
 26

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 707 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein  
 (B) LOCATION: 1..707  
 (D) OTHER INFORMATION: /note= "EXTRACELLULAR DOMAIN OF  
 PSMA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Lys Ser Ser Asn Glu Ala Thr Asn Ile Thr Pro Lys His Asn Met Lys  
 1 5 10 15  
 Ala Phe Leu Asp Glu Leu Lys Ala Glu Asn Ile Lys Lys Phe Leu Tyr  
 20 25 30  
 Asn Phe Thr Gln Ile Pro His Leu Ala Gly Thr Glu Gln Asn Phe Gln  
 35 40 45  
 Leu Ala Lys Gln Ile Gln Ser Gln Trp Lys Glu Phe Gly Leu Asp Ser  
 50 55 60  
 Val Glu Leu Ala His Tyr Asp Val Leu Leu Ser Tyr Pro Asn Lys Thr  
 65 70 75 80  
 His Pro Asn Tyr Ile Ser Ile Ile Asn Glu Asp Gly Asn Glu Ile Phe  
 85 90 95  
 Asn Thr Ser Leu Phe Glu Pro Pro Pro Gly Tyr Glu Asn Val Ser  
 100 105 110  
 Asp Ile Val Pro Pro Phe Ser Ala Phe Ser Pro Gln Gly Met Pro Glu  
 115 120 125  
 Gly Asp Leu Val Tyr Val Asn Tyr Ala Arg Thr Glu Asp Phe Phe Lys  
 130 135 140  
 Leu Glu Arg Asp Met Lys Ile Asn Cys Ser Gly Lys Il Val Ile Ala  
 145 150 155 160

- 41 -

Arg Tyr Gly Lys Val Phe Arg Gly Asn Lys Val Lys Asn Ala Gln Leu  
 165 170 175  
 5 Ala Gly Ala Lys Gly Val Ile Leu Tyr Ser Asp Pro Ala Asp Tyr Phe  
 180 185 190  
 Ala Pro Gly Val Lys Ser Tyr Pro Asp Gly Trp Asn Leu Pro Gly Gly  
 195 200 205  
 10 Gly Val Gln Arg Gly Asn Ile Leu Asn Leu Asn Gly Ala Gly Asp Pro  
 210 215 220  
 Leu Thr Pro Gly Tyr Pro Ala Asn Glu Tyr Ala Tyr Arg Arg Gly Ile  
 225 230 235 240  
 15 Ala Glu Ala Val Gly Leu Pro Ser Ile Pro Val His Pro Ile Gly Tyr  
 245 250 255  
 Tyr Asp Ala Gln Lys Leu Leu Glu Lys Met Gly Gly Ser Ala Pro Pro  
 260 265 270  
 20 Asp Ser Ser Trp Arg Gly Ser Leu Lys Val Pro Tyr Asn Val Gly Pro  
 275 280 285  
 25 Gly Phe Thr Gly Asn Phe Ser Thr Gln Lys Val Lys Met His Ile His  
 290 295 300  
 Ser Thr Asn Glu Val Thr Arg Ile Tyr Asn Val Ile Gly Thr Leu Arg  
 305 310 315 320  
 30 Gly Ala Val Glu Pro Asp Arg Tyr Val Ile Leu Gly Gly His Arg Asp  
 325 330 335  
 Ser Trp Val Phe Gly Gly Ile Asp Pro Gln Ser Gly Ala Ala Val Val  
 340 345 350  
 35 His Glu Ile Val Arg Ser Phe Gly Thr Leu Lys Lys Glu Gly Trp Arg  
 355 360 365  
 40 Pro Arg Arg Thr Ile Leu Phe Ala Ser Trp Asp Ala Glu Glu Phe Gly  
 370 375 380  
 Leu Leu Gly Ser Thr Glu Trp Ala Glu Glu Asn Ser Arg Leu Leu Gln  
 385 390 395 400  
 45 Glu Arg Gly Val Ala Tyr Ile Asn Ala Asp Ser Ser Ile Glu Gly Asn  
 405 410 415  
 Tyr Thr Leu Arg Val Asp Cys Thr Pro Leu Met Tyr Ser Leu Val His  
 420 425 430  
 50 Asn Leu Thr Lys Glu Leu Lys Ser Pro Asp Glu Gly Phe Glu Gly Lys  
 435 440 445  
 Ser Leu Tyr Glu Ser Trp Thr Lys Lys Ser Pro Ser Pro Glu Phe Ser  
 450 455 460  
 Gly Met Pro Arg Ile Ser Lys Leu Gly Ser Gly Asn Asp Phe Glu Val  
 465 470 475 480  
 60 Phe Phe Gln Arg Leu Gly Ile Ala Ser Gly Arg Ala Arg Tyr Thr Lys  
 485 490 495  
 Asn Trp Glu Thr Asn Lys Phe Ser Gly Tyr Pro Leu Tyr His Ser Val  
 500 505 510  
 65

- 42 -

Tyr Glu Thr Tyr Glu Leu Val Glu Lys Phe Tyr Asp Pro Met Phe Lys  
 515 520 525  
 5 Tyr His Leu Thr Val Ala Gln Val Arg Gly Gly Met Val Phe Glu Leu  
 530 535 540  
 Ala Asn Ser Ile Val Leu Pro Phe Asp Cys Arg Asp Tyr Ala Val Val  
 545 550 555 560  
 10 Leu Arg Lys Tyr Ala Asp Lys Ile Tyr Ser Ile Ser Met Lys His Pro  
 565 570 575  
 15 Gln Glu Met Lys Thr Tyr Ser Val Ser Phe Asp Ser Leu Phe Ser Ala  
 580 585 590  
 Val Lys Asn Phe Thr Glu Ile Ala Ser Lys Phe Ser Glu Arg Leu Gln  
 595 600 605  
 20 Asp Phe Asp Lys Ser Asn Pro Ile Val Leu Arg Met Met Asn Asp Gln  
 610 615 620  
 Leu Met Phe Leu Glu Arg Ala Phe Ile Asp Pro Leu Gly Leu Pro Asp  
 625 630 635 640  
 25 Arg Pro Phe Tyr Arg His Val Ile Tyr Ala Pro Ser Ser His Asn Lys  
 645 650 655  
 30 Tyr Ala Gly Glu Ser Phe Pro Gly Ile Tyr Asp Ala Leu Phe Asp Ile  
 660 665 670  
 Glu Ser Lys Val Asp Pro Ser Lys Ala Trp Gly Glu Val Lys Arg Gln  
 675 680 685  
 35 Ile Tyr Val Ala Ala Phe Thr Val Gln Ala Ala Ala Glu Thr Leu Ser  
 690 695 700  
 Glu Val Ala  
 705

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

55 CTTAAGCGAG CCCAGAGGGC CCACA  
 25

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- 43 -

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTCGAGCTCA TTTACCCGGA GTCCG  
25

10

- 44 -

What is claimed is:

- 1 1. A DNA produced by recombinant DNA techniques for inducing expression  
2 and subsequent secretion of a target protein, said sequence being free of  
3 immunoglobulin CH1 and comprising a polynucleotide encoding, from its 5' to  
4 3' direction:
  - 5 A) a secretion cassette which comprises  
6 a signal sequence;  
7 an immunoglobulin Fc region; and  
8 B) a target protein sequence.
- 1 2. The DNA of claim 1 wherein the target protein sequence encodes a  
2 soluble form of prostate specific membrane antigen.
- 1 3. The DNA of claim 1 wherein the target protein sequence encodes at  
2 least a portion of gp120 protein.
- 1 4. The DNA of claims 1, 2 or 3 wherein the signal sequence encodes a  
2 signal peptide which directs secretion of the target protein and is then  
3 removed by enzymatic cleavage.
- 1 5. The DNA of claim 1, 2 or 3 wherein the Fc region is altered to delete  
2 at least one effector function activity.
- 1 6. The DNA of claim 1, 2 or 3 wherein the Fc region comprises a hinge, a  
2 CH2 domain and a CH3 domain of immunoglobulin gamma.
- 1 7. The DNA of claim 1, 2 or 3 wherein the Fc region comprises a hinge  
2 region and a CH3 domain of immunoglobulin gamma.
- 1 8. The DNA of claim 1, 2 or 3 further comprising a proteolytic cleavage  
2 site interposed 3' of a portion of said polynucleotide encoding said  
3 immunoglobulin Fc region and 5' of a portion of said polynucleotide encoding  
4 said entire target protein.
- 1 9. A replicable expression vector for transfecting a mammalian cell, said  
2 vector comprising the DNA of claim 1, 2 or 3.

- 45 -

- 1 10. A host cell transformed with the DNA of claim 1, 2 or 3.
- 1 11. Target protein produced by culturing the cell of claim 10.
- 1 12. A method of producing a target protein comprising the steps of:  
2 1) transfecting the DNA of claim 1, 2 or 3 into a host cell;  
3 2) culturing the host cell in a medium under conditions to promote  
4 expression and secretion of a fusion protein comprising  
5 an immunoglobulin Fc region, and  
6 a target protein; and  
7 3) collecting the fusion protein from the medium.
- 1 13. The method of claim 12 wherein the fusion protein has bioactivity of  
2 the target protein.
- 1 14. The method of claim 12 further comprising the additional steps, after  
2 step 3, of  
3 4) cleaving the Fc region from the target protein, and  
4 5) collecting the target protein.
- 1 15. A method of producing a target protein comprising the steps of:  
2 1) transfecting the DNA of claim 1, 2 or 3 into a host cell;  
3 2) culturing the host cell in a medium under conditions to promote  
4 expression and secretion of a fusion protein comprising  
5 an immunoglobulin Fc region, and  
6 a target protein; and  
7 3) cleaving the Fc region from the target protein, and  
8 4) collecting the target protein from the medium.

1/1

FIG. 1a

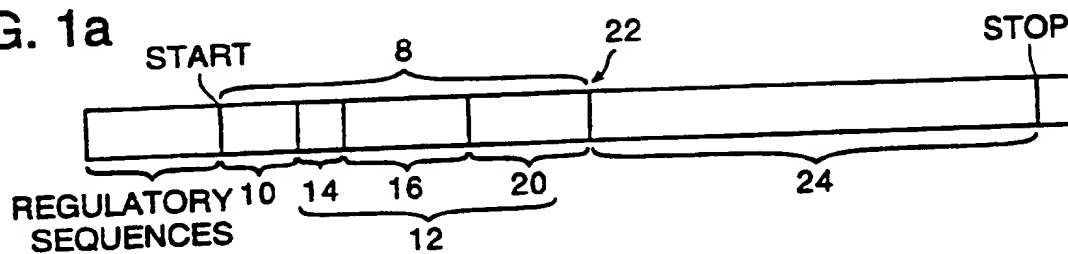


FIG. 1b

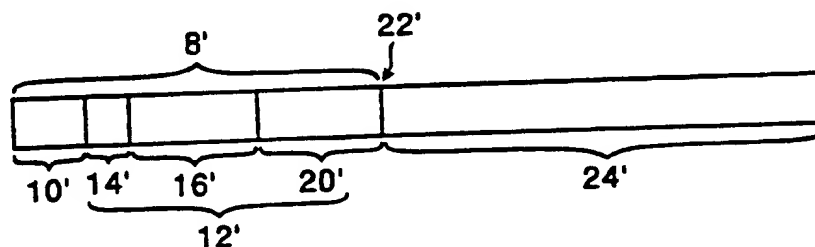


FIG. 1c

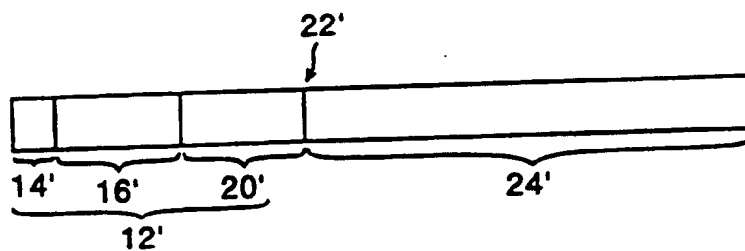
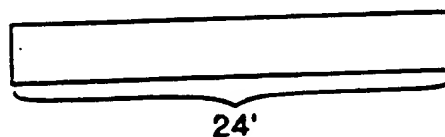


FIG. 1d





# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 95/11720

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 C12N15/62 C12N15/85 C12N5/10 C07K14/705 C07K14/16  
C07K16/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BIOCONJUGATE CHEMISTRY, vol. 4, 1993 AMERICAN CHEM. SOC., US, pages 230-235, S.D. GILLIES ET AL. 'Biological activity and in vivo clearance of antitumor antibody/cytokine fusion protein' cited in the application see the whole document ---	2-15
Y	PROC. NATL. ACAD. SCI., vol. 89, February 1992 NATL. ACAD SCI., WASHINGTON, DC, US;, pages 1428-1432, S.D. GILLIES ET AL. 'Antibody-targeted interleukin 2 stimulates T-cell killing of autologous tumor cells' cited in the application see the whole document ---	2-15

-/-

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*A\* document member of the same patent family

Date of the actual completion of the international search

22 January 1996

Date of mailing of the international search report

28.02.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Hornig, H

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US 95/11720

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,91 00360 (MEDAREX INC) 10 January 1991 see page 2, line 26 - page 12, line 23 ---	2-15
Y	WO,A,91 08298 (GENENTECH INC) 13 June 1991 see page 4, line 29 - line 32 see page 8, line 28 - line 30 see page 10, line 29 - line 31 see page 10, line 34 - line 37 ---	2-15
Y	WO,A,92 08495 (ABBOTT BIOTECH INC) 29 May 1992 see page 10, line 2 - page 14, line 23 see page 24, line 1 - line 33 ---	2-15
Y	WO,A,92 08801 (ABBOTT LAB) 29 May 1992 see page 7, line 15 - page 16, line 13 ---	2-15
Y	EP,A,0 511 747 (ROHM & HAAS) 4 November 1992 see page 4, line 24 - line 30 ---	4,14,15
Y	WO,A,93 03157 (DANA FARBER CANCER INST INC) 18 February 1993 see page 14, line 6 - line 23 ---	4,14,15
A	PROC. NATL.ACAD SCI., vol. 91, no. 3, 1 February 1994 NATL. ACAD SCI., WASHINGTON, DC, US;, pages 989-993, H.-G. CHEON ET AL. 'High-affinity binding sites for related fibroblast growth factor ligands reside within different receptor immunoglobulin-like domain' see the whole document ---	1-15
A	HUMAN ANTIBODIES AND HYBRIDOMAS, vol. 1, no. 1, 1990 BUTTERWORTH PUBLISHERS, STONEHAM, MA, US, pages 47-54, S.D. GILLIES AND J.S. WESOLOWSKI 'Antigen binding and biological activities of engineered mutant chimeric antibodies with human tumor specificities' cited in the application see the whole document ---	1-15
A	NATURE, vol. 337, 9 February 1989 MACMILLAN JOURNALS LTD., LONDON, UK, pages 525-531, D.J. CAPON ET AL. 'Designing CD4 immunoadhesins for AIDS therapy' see the whole document ---	1-15
	---	

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 95/11720

## C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. IMMUNOL. METHODS, vol. 125, 1989 ELSEVIER, AMSTERDAM NL, pages 191-202, S.D. GILLIES ET AL. 'High-level expression of chimeric antibodies using adapted cDNA variable region cassettes' cited in the application see the whole document -----	1-15
P,X	WO,A,95 21258 (US GOVERNMENT) 10 August 1995 see page 7, line 9 - page 11, line 18 see page 14, line 15 - page 15, line 37 see page 16, line 1 - page 17, line 19 -----	1

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US 95/11720

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9100360	10-01-91	AU-B- 641673 AU-B- 5962290 CA-A- 2062795 EP-A- 0479909	30-09-93 17-01-91 30-12-90 15-04-92
WO-A-9108298	13-06-91	EP-A- 0526452 JP-T- 5503009	10-02-93 27-05-93
WO-A-9208495	29-05-92	AU-B- 660297 AU-B- 9059691 CA-A- 2095836 EP-A- 0574395	22-06-95 11-06-92 10-05-92 22-12-93
WO-A-9208801	29-05-92	CA-A- 2095842 EP-A- 0556328	10-05-92 25-08-93
EP-A-0511747	04-11-92	AU-B- 659139 AU-B- 1398792 JP-A- 6166698 NZ-A- 242395	11-05-95 22-10-92 14-06-94 23-12-93
WO-A-9303157	18-02-93	NONE	
WO-A-9521258	10-08-95	AU-B- 1690295	21-08-95